Bert Stoneberg - Ravitch, New York Times, November 7, 2005

From: Bert Stoneberg
To: communications@brookings.edu
Date: 11/8/2005 12:07 PM
Subject: Ravitch, New York Times, November 7, 2005

Dr. Diane Ravitch,

I know you personally believe that "the states have embraced low standards and grade inflation" (New York Times, November 7, 2005). There was, however, nothing in the NAEP mathematics data you cited from Idaho (nor other states) that supports your opinion.

The confirmatory evidence should not have to reach such a high level standard as “beyond a reasonable doubt.” Instead, states should be given the benefit of the doubt about whether their results are confirmed. […]The state results may be questioned when there is consistent, compelling contrary evidence from the National Assessment that cannot be explained simply by the differences between the two tests or other relevant factors. (Ad Hoc Committee on Confirming Test Results, p.9)

Two relevant factors your opinion piece ignored are (1) differences in the NAEP and the states definitions of "basic" and "proficient", and (2) NAEP sample sizes required to make valid inferences. As a consequence, you incorrectly compared the percentage making AYP (i.e., at or above state proficient) with the percentage at or above NAEP *Proficient*. The NAEP statistic most comparable with a state's percentage of students meeting AYP is the percent at or above NAEP *Basic*.

You know from experience that NAGB does not make technical policy decisions on a whim. Following up on the Mosquin and Chromy report (see excerpts and reference below), NAGB modified its policy to facilitate using NAEP to confirm state AYP results as put forth in NCLB. Starting with 2005 reports an intentional focus on the proportion of students "at or above NAEP *Basic*" was reported. To witness the update in NAGB policy, you need only examine the State Snapshots Reports showing fourth grade mathematics performance in 2003 and 2005 that NCES/ETS prepared for the Idaho. Look at the sliding bar charts in the upper right hand section of the snapshots. Idaho's 2003 and 2005 snapshots are available on the web at

2003: http://www.sde.state.id.us/naep/data/ma03/ma03-04-snapshot.pdf
2005: http://www.sde.state.id.us/naep/data/ma05/ma05-nces-snapshot-ID4.pdf

The "price of this local watering down" of fourth grade mathematics in Idaho is clear, if you care to take a look. Even using the lofty criterion of "at or above NAEP *Proficient*", we claim that 20 percent met the criterion in 2000, 31 percent met it in 2003, and 40 percent in 2005. With "at or above NAEP *Basic*", we claim 68 percent for 2000, 80 percent for 2003, and 86 percent for 2005. And, by the way, the 86 percent at or above NAEP *Basic* in 2005 doesn't seem all that far off from the 90 percent at or above Idaho proficient reported for AYP in 2005, especially considering all

the differences between the two tests.

Technical Excepts from Mosquin and Chromy

In selecting a gap performance measure, comparability with the AYP statistic is more important than correlation. Adequate yearly progress is already defined within the Act based on the percentage of scores exceeding the basic proficiency level. The basic proficiency level corresponds roughly to the percentage *below basic* on the NAEP scale. Therefore, of the various statistics that might be used for measuring a gap on the NAEP scale—proportion at or above the basic, proficient, or advanced achievement level, or mean standardized score—the proportion at or above the basic achievement level will both have the greatest correlation with the adequate yearly progress statistic and also be the most directly comparable. (Mosquin and Chromy, p.12)

Using the proportion of students at or above the basic achievement level required sample sizes larger than using the mean scale score, but smaller than when using the proportion of students at or above the proficient level. Thus the proportion of students at or above the basic achievement level appears to be the more usable of the two achievement level performance measures. It also appears to be most compatible with the AYP statistic, providing a consistent quantitative measure for both gaps and adequate yearly progress. (Mosquin and Chromy, p. 41)

References

Ad Hoc Committee on Confirming Test Results. (2002). *Using the National Assessment of Educational Progress to Confirm State Test Results*. Washington, D.C.: U.S. Department of Education, National Assessment Governing Board.

Mosquin, P., and Chromy J. (2004). *Federal Sample Sizes for Confirmation of State Tests in the No Child Left Behind Act*. Washington, D.C.: American Institutes for Research, NAEP Validity Studies Panel.

Bert Stoneberg
NAEP State Coordinator
State Department of Education
Boise, Idaho
Visit http://www.sde.state.id.us/naep/

Please Don't Use NAEP Scores to Rank Order the 50 States
Available online: http://pareonline.net/getvn.asp?v=10&n=9

Everything should be as simple as possible,
but not simpler. - Albert Einstein

Every State Left Behind

By DIANE RAVITCH
for the New York Times
Published: November 7, 2005

WHILE in office, Presidents George H. W. Bush and Bill Clinton both called for national academic standards and national tests in the public schools. In both cases, the proposals were rejected by a Congress dominated by the opposing party. The current President Bush, with a friendly Congress in hand, did not pursue that goal because it is contrary to the Republican Party philosophy of localism. Instead he adopted a strategy of "50 states, 50 standards, 50 tests" - and the evidence is growing that this approach has not improved student achievement. Americans must recognize that we need national standards, national tests and a national curriculum.

The release last month of test results by the National Assessment of Educational Progress, which is part of the Department of Education, vividly demonstrated why varying state standards and tests are inadequate. Almost all states report that, based on their own tests, incredibly large proportions of their students meet high standards. Yet the scores on the federal test (which was given to a representative sample of fourth and eighth graders) were far lower. Basically, the states have embraced low standards and grade inflation.

Idaho claims that 90 percent of its fourth-grade students are proficient in mathematics, but on the federal test only 41 percent reached the Education Department's standard of proficiency. Similarly, New York reports that nearly 85 percent of its fourth graders meet state standards in mathematics, yet only 36 percent tested as proficient on the national assessment. North Carolina boasts an impressive 92 percent pass rate on the state test, but only 40 percent meet the federal standard.

In fourth-grade reading, the gaps between state and national reports are equally large. Georgia claims that 87 percent of its pupils are proficient in reading, but only 26 percent reached that level on the national exam. Alabama says that 83 percent of its students are proficient, but only 22 percent meet the federal standard.

The same discrepancies are found in the scores for eighth-grade reading, where Texas reports that 83 percent met the state standard, but the federal test finds that only 26 percent are proficient. Tennessee and North Carolina both claim that 88 percent are proficient readers, whereas 26 percent and 27 percent, respectively, met that mark on the federal test.

Why the discrepancies? The states function in a political environment. Educational leaders and elected officials want to assure the public that the schools are doing their jobs and making progress. The federal testing program, administered for the past 15 years by an independent, bipartisan governing board, has never been cowed by the demands of parents, school officials and taxpayers for good news.

In the No Child Left Behind law of 2001, Congress left it to each state to develop its own standards and tests, but added that the tests given by National Assessment of Educational Progress should serve as an external gauge of national and state-level achievement. The federal tests are considered the gold standard for good reason: they are the product of a long-term federal investment in research and development. Unlike the state tests, the federal program tries to align its

performance standards with international education standards. Many states model their testing on the national program, but still cling to lower standards for fear of alienating the public and embarrassing public officials responsible for education.

The price of this local watering-down is clear. Our fourth-grade students generally do well when compared with their peers in other nations, but eighth-grade students are only average globally, and 12th graders score near the bottom in comparison with students in many European and Asian nations. Even our students who have taken advanced courses in mathematics and physics perform poorly relative to their peers on international tests.

Last month, the National Academy of Sciences released a report warning that our nation's "strategic and economic security," as well as our leadership in the development of new technologies, is at risk unless we invest heavily in our human capital; that is, the education of our people. The academy report made clear that many young Americans do not know enough about science, technology or mathematics to understand or contribute to the evolving knowledge-based society. The best way to compete in the global economy, the report maintained, is to ensure that American workers are "the best educated, the hardest-working, best trained, and most productive in the world."

It is fair to say that we will not reach that goal if we accept mediocre performance and label it "proficient." Nor will we reach that goal if we pretend that mathematics taught in Alaska or Iowa is profoundly different from the mathematics taught in Maine or Florida, or for that matter, in Japan and Hungary.

Unfortunately, the political calculations that resulted in the No Child Left Behind law adopting a strategy of letting the states choose their own standards and tests remain the reality. In general, Republicans are wary of national standards and a national curriculum, while Democrats are wary of testing in general. Both parties must come to understand that the states are not competing with each other to ratchet up student achievement. Instead, they are maintaining standards that meet the public's comfort level.

America will not begin to meet the challenge of developing the potential of our students until we have accurate reporting about their educational progress. We will not have accurate reporting until that function is removed from the constraints of state and local politics. We will be stuck with piecemeal and ineffective reforms until we agree as a nation that education - not only in reading and mathematics, but also science, history, literature, foreign languages and the arts - must be our highest domestic priority.

Diane Ravitch, a research professor at New York University and fellow at the Brookings Institution, was on the governing board of the National Assessment of Educational Progress from 1997 to 2004.